Directional derivatives of optimal solutions in smooth nonlinear programming

J. Frederic Bonnans

UNITÉ DE RECHERCHE
INRIA-ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél.:(1) 39 63 55 11

Rapports de Recherche

N° 1006

Programme 5

DIRECTIONAL DERIVATIVES OF OPTIMAL SOLUTIONS IN SMOOTH NONLINEAR PROGRAMMING

J. Frédéric BONNANS

Mars 1989

DIRECTIONAL DERIVATIVES OF OPTIMAL SOLUTIONS IN SMOOTH NONLINEAR PROGRAMMING

J. Frédéric BONNANS
INRIA - BP 105
78153 Rocquencourt

Abstract

We consider a smooth nonlinear program subject to perturbations in the right-hand-side of the constraints. We do not assume that the unique solution of the original problem satisfies any qualification hypothesis. We suppose instead that the direction of perturbation satisfies the hypothesis of Gollan. We study the variation of the cost and with the help of some second-order sufficiency conditions obtain some conditions satisfied by the first term of the expansion of the solution. These conditions may vary depending on the existence of a Lagrange multiplier for the original problem.

DERIVEES DIRECTIONNELLES DE SOLUTIONS DE PROGRAMMES NON LINEAIRES LISSES

Résumé

Soit un programme nonlinéaire lisse soumis à des perturbations dans le membre de droite des contraintes. Nous ne supposons vérifiée aucune hypothèse de qualification pour le problème non perturbé. A la place nous supposons que la direction de perturbation satisfait l'hypothèse de Gollan. Nous étudions la variation du coût et, grâce à certaines conditions suffisantes du deuxième ordre, obtenons des conditions satisfaites par le premier terme du développement des solutions. Ces conditions varient suivant l'existence d'un multiplicateur de Lagrange pour le problème original.

Programme 5

DIRECTIONAL DERIVATIVES OF OPTIMAL SOLUTIONS IN SMOOTH NONLINEAR PROGRAMMING*

Joseph Frédéric BONNANS
INRIA-Rocquencourt - BP 105
78153 Le Chesnay - France

March 9, 1989

Abstract

We consider a smooth nonlinear program subject to perturbations in the right-hand-side of the constraints. We do not assume that the unique solution of the original problem satisfies any qualification hypothesis. We suppose instead that the direction of perturbation satisfies the hypothesis of Gollan. We study the variation of the cost and with the help of some second-order sufficiency conditions obtain some conditions satisfied by the first term of the expansion of the solution. These conditions may vary depending on the existence of a Lagrange multiplier for the original problem.

1 INTRODUCTION

This paper is concerned with the perturbation of a standard nonlinear program with smooth data. The first results in this direction where obtained by applying the implicit function theorem (see Fiacco [3] for a review of these results). Then the hypothesis of linear independence of the gradients of active constraints, strict complementary and second order sufficiency conditions were needed. Jittorntrum [8] relaxed the strict complementarity hypothesis. With the help of a strong second-order sufficiency condition he obtained the directional differentiability of a local solution. Using a semi-strong second-order sufficiency condition the author generalized this result

*Presented at the workshop "Analyse quantitative de la sensibilité en optimisation", S^{te} Adèle (Québec), Feb. 19-23, 1989.

by studying in [1] situations where a finite number of computable directions contain the limit points of the ratio of solutions w.r.t. perturbation. In [1] an algorithm is given that allows to compute effectively the solutions of the perturbed problem. Recently Shapiro [11] using the Mangasarian and Fromovitz qualification hypothesis [10] and some second-order sufficiency conditions gave a second-order analysis of the optimal value and formulated an extended quadratic problem for which directional derivatives of the solutions are optimal. On the other hand, following an idea of Gollan [7], Gauvin and Janin studied in[4] [5] [6] the directional derivative of the value functions and the magnitude of the variation in the solution.

Our aim in this paper is to improve these results in two directions. First when the original problem has no Lagrange multiplier (a situation that has not been much considered) we give an estimate of the variation of the optimal cost (which is of the order of the square root of the perturbation) and formulate a subproblem with quadratic constraints whose solutions have a close connection with the variations in the solution. When some Lagrange multiplier exists we obtain some relations concerning the second-order analysis of the cost and the first-order variation of the solutions ; these results improve slightly those of Shapiro [11] in the case when the hypothesis of Mangasarian and Fromovitz is also satisfied.

The paper is organized as follows. In part 2 we give some notations and preliminary results. Part 3 (resp. part 4) is concerned with the case when no Lagrange multiplier exists (resp. there exist some Lagrange multipliers). Part 5 presents two examples in which the results of part 3 and part 4 allow to compute effectively the first term of the expansion of the solutions.

2 PRELIMINARIES

This paper is concerned with the following problem :

$$P(u) \qquad \min f(x); x \in \mathbb{R}^n; g_I(x) \leq u_I, g_J(x) = u_J$$

with $I = (1, \ldots, q), J = (q+1, \ldots, p), u \in \mathbb{R}^p, f$ and g being three times continuously differentiable. In order to shorten the notations, for y, z in $\mathbb{R}^p$ we will write $y << z$ whenever $y_I \leq z_I$ and $y_J = z_J$. More generally for K such that $K \subset I$ we will write

$$y \overset{K}{<<} z$$

whenever

$$\begin{cases} y_i & \leq & z_i, i \in K, \\ y_J & = & z_J. \end{cases}$$

The set of feasible points of $P(u)$ is

$$F(u) := \{x \in I\!R^n ; g(x) << u\},$$

and the set of solutions of $P(u)$ is

$$S(u) := \{x \in F(u); f(x) = \inf P(u)\}.$$

For any x in $I\!R^n$ we will say that $\lambda = (\lambda_0, \ldots, \lambda_p) \in I\!R^{p+1}$ is a multiplier associated to x for problem $P(u)$ if $\lambda \neq 0$ and (x, λ) satisfies the following first order optimality conditions

$$\begin{cases} \lambda_0 \nabla f(x) & + \sum_{i=1}^{p} \lambda_i \nabla g_i(x) = 0, \\ g(x) << u \quad , & \lambda_0 \geq 0, \lambda_I \geq 0, \lambda_i(g_i(x) - u_i) = 0, \forall i \in I. \end{cases}$$

It is well known that to each local solution of $P(u)$ is associated at least one multiplier. We denote by $\Lambda(u, x)$ the set of multipliers. Subsets of $\Lambda(u, x)$ of interest are the set of degenerate multipliers

$$\Lambda_0(u, x) = \{\lambda \in \Lambda(u, x); \lambda_0 = 0\}$$

and of Lagrange multipliers

$$\Lambda_1(u, x) = \{\lambda \in \Lambda(u, x); \lambda_0 = 1\}.$$

These degenerate multipliers and Lagrange multipliers can often be identified without confusion to elements of $I\!R^p$; e.g. for z in $I\!R^p$, by $\lambda^t z$ we will mean $\sum_{i=1}^{p} \lambda_i u_i$. We will denote $\Lambda_i(0, \bar{x})$ by Λ_i $(i = 1, 2)$.

In all this paper we will assume that the two following hypotheses hold :

$$(H1) \qquad S(0) = \{\bar{x}\}$$

(i.e. $P(0)$ has a unique solution) and an inf-compactness condition :

$$(H2) \qquad \exists \alpha > 0, \beta > 0 : \; if \|u\| \leq \beta \; and \; x \in F(u) \; then \; \|x\| \leq \alpha.$$

Let us define the set of active constraints

$$I(u, x) = \{i \in I; g_i(x) = u\}.$$

We will denote $I(0, \bar{x})$ by $\bar{I}$.

We will say that a point x in $F(u)$ satisfies the MF (Mangasarian and Fromovitz) hypothesis for some subset $K \subset I$ if the following holds

(i) $\{\nabla g_i(x)\}_{i \in J}$ is linearly independent
(ii) there exists d in $\ker g'_J(x)$ satisfying $g'_K(x)\, d < 0$

If this relation is satisfied for $K = I(u, x)$ and $x \in F(u)$ we will say that x satisfies MF for problem $P(u)$ (this is the classical hypothesis of Mangasarian and Fromovitz [10]).

The case when $\bar{x}$ satisfies MF for problem P(0) has been studied in detail by Shapiro [11]. We will relax this hypothesis. However in order to ensure the non-emptiness of the feasible set of the perturbed problem we will assume that the following holds :

$$(H3) \qquad \lambda^t u > 0 \text{ for all } \lambda \text{ in } \Lambda_0.$$

A condition equivalent to $H3$ is given below.

Lemma 2.1 *$H3$ is equivalent to the following requirement :*

$$\begin{array}{ll} (i) & \{\nabla g_i(\bar{x})\}_{i \in J} \text{ is linearly independent.} \\ (ii) & \exists \bar{d} \text{ in } \mathbb{R}^n \text{ and } \gamma > 0 \text{ such that} \\ & g'_J(\bar{x})\bar{d} = u_J \\ & g'_i(\bar{x})\bar{d} + \gamma \leq u_i \quad \text{for all } i \text{ in } \bar{I}. \end{array}$$

Proof Hypothesis $H3$ says that the following system has no non-null solution

$$\left\{ \begin{array}{lcl} \sum_{i=1}^{p} \lambda_i \nabla g_i(\bar{x}) & = & 0, \\ \sum_{i=1}^{p} \lambda_i u_i & \leq & 0, \\ \lambda_I \geq 0, \lambda_i & = & 0 \text{ if } g_i(\bar{x}) < 0, \forall i \in I. \end{array} \right.$$

This is equivalent to say that the following system has no non-null solution (here $\theta \in \mathbb{R}$) :

$$\sum_{i=1}^{p} \lambda_i \begin{pmatrix} \nabla g_i(\bar{x}) \\ -u_i \end{pmatrix} + \theta \begin{pmatrix} 0 \\ -1 \end{pmatrix} = 0,$$
$$\lambda_I \geq 0, \theta \geq 0; \lambda_i = 0 \text{ if } g_i(\bar{x}) < 0, \forall i \in I.$$

But this is the dual of the MF condition applied to the system (here $\sigma \in \mathbb{R}$)

$$\left\{ \begin{array}{lcl} g'(\bar{x})d - \sigma u & \stackrel{I}{<<} & 0, \\ -\sigma & \leq & 0. \end{array} \right.$$

Hence $H3$ is equivalent to

$$\left\{ \begin{pmatrix} \nabla g_i(\bar{x}) \\ -u_i \end{pmatrix} \right\}_{i \in J} \text{ linearly independent}$$

i.e. $(d, \sigma) \rightarrow g'_J(\bar{x})d - \sigma u_J$ surjective and there exists $(\bar{d}, \bar{\sigma})$ in $\mathbb{R}^n \times \mathbb{R}$ satisfying

$$\begin{cases} g'_J(\bar{x})\bar{d} - \bar{\sigma}u_J &= 0, \\ g'_i(\bar{x})\bar{d} - \bar{\sigma}u_i &< 0 \text{ for all } i \text{ in } \bar{I}, \\ -\bar{\sigma} &< 0. \end{cases}$$

These above conditions are an easy consequence of (i), (ii), hence we just have to prove that they imply (i) (ii). As $\bar{\sigma} > 0$, multiplying $(\bar{d}, \bar{\sigma})$ by $\frac{1}{\bar{\sigma}}$ we find (ii). The first condition says that for any e in $I\!R^p$ there exists (d, σ) such that

$$g'_J(\bar{x})d - \sigma u_J = e$$

From this and $g'_J(\bar{x})\bar{d} - \bar{\sigma}u_J = 0$ we deduce that

$$g'_J(\bar{x})(d - (\sigma/\bar{\sigma})\bar{d}) = e$$

hence $g'_J(\bar{x})$ is surjective, i.e. (i) holds. $\square$

From Lemma 2.1, we will deduce a nice property concerning $F(\sigma u)$ when u satisfies $H3$. We define the two following closed sets :

$$\tau = \left\{ d \in I\!R^n; \exists \sigma^k \searrow 0, x^k \in F(\sigma^k u); \frac{x^k - \bar{x}}{\sigma^k} \to d \right\},$$

$$T = \left\{ d \in I\!R^n; g'(\bar{x})d \overset{I}{<<} u \right\}.$$

Lemma 2.2 *Hypothesis $H3$ implies that $\tau = T$. In addition for any d in T there exists $x(\sigma) = \bar{x} + \sigma d + 0(\sigma^2)$ in $F(\sigma u)$ for $\sigma > 0$ small enough.*

<u>Proof</u> The inclusion $\tau \subset T$ is easily obtained. Conversely, pick d in T. For $\beta > 0$ given consider

$$y(\sigma) = \bar{x} + \sigma \left[(1 - \beta\sigma)d + \beta\sigma\bar{d} \right],$$

where $\bar{d}$ satisfies the conditions of Lemma 2.1. Then for i in $\bar{I}$:

$$\begin{aligned} g_i(y(\sigma)) &= \sigma g'_i(\bar{x}) \left[(1 - \beta\sigma)d + \beta\sigma\bar{d} \right] + 0(\sigma^2), \\ &\leq \sigma(1 - \beta\sigma)u_i + \beta\sigma^2(u_i - \gamma) + 0(\sigma^2) \\ &= \sigma u_i - \beta\gamma\sigma^2 + 0(\sigma^2). \end{aligned}$$

Here when $\beta\sigma < 1$ we have $\|(1 - \beta\sigma)d + \beta\sigma\bar{d}\| \leq \max(\|d\|, \|\bar{d}\|)$ hence we may suppose that if $\sigma < 1/\beta$ the term $0(\sigma^2)$ does not depend on β. Taking $\beta > 0$ large enough we obtain when $\beta\sigma < 1$:

$$g_i(y(\sigma)) \leq \sigma(u_i - \sigma\beta\gamma/2).$$

Similarly for i in J and $\beta\sigma < 1$ we have

$$g_i(y(\sigma)) = \sigma u_i + 0(\sigma^2).$$

Hence (from the implicit function theorem) there exists $r(\sigma) = 0(\sigma^2)$ such that $x(\sigma) = y(\sigma) + r(\sigma)$ satisfies $g_J(x(\sigma)) = \sigma u_J$, and for β large enough and σ small we will still obtain

$$g_i(x(\sigma)) = \sigma(u_i - \sigma\beta\gamma/2), i \in \bar{I}$$

Hence $g_i(x(\sigma)) \leq \sigma u_i, i \in \bar{I}$, i.e. $x(\sigma) = \bar{x} + \sigma d + 0(\sigma^2)$ is in $F(\sigma u)$ as desired. □

Remark 2.1 If $H3$ holds, using $\bar{d}$ we can construct a path $x(\sigma)$ with $x(\sigma) \in F(\sigma u)$. This proves in particular that $F(\sigma u) \neq \emptyset$.

3 THE CASE $\Lambda_1 = \emptyset$

In this part we present a study of the perturbation problem when $\Lambda_1 = \emptyset$. A first result in this direction is

Theorem 3.1 *If hypotheses $(H1), (H2), (H3)$ hold and $\Lambda_1 = \emptyset$, then*

$$\limsup_{\sigma \searrow 0} \frac{\inf P(\sigma u) - f(\bar{x})}{\sigma^{1/2}} < 0.$$

Proof As in Lemma 2.2. we construct a feasible path, but on a different basis. The homogeneous linear program

$$\min_{d \in R^n} \nabla f(\bar{x})^t d; g'(\bar{x})d \overset{\bar{I}}{<<} 0$$

has $d = 0$ as a feasible point but its infimum is $-\infty$ (otherwise $d = 0$ would be a solution and Λ_1 would not be empty). Consequently there exists d^0 satisfying

$$\nabla f(\bar{x})^t d^0 = -1; g'(\bar{x})d^0 \overset{\bar{I}}{<<} 0.$$

For α given in $]0, 1]$, we consider the path

$$y(\sigma) = \bar{x} + \alpha\sigma^{1/2}d^0 + \sigma\bar{d}.$$

Here $\bar{d}$ is given by Lemma 2.1. As $|\alpha| \leq 1$ we have

$$
\begin{aligned}
g_J(y(\sigma)) &= g_J'(\bar{x})\left(\alpha\sigma^{1/2}d^0 + \sigma\bar{d}\right) + \alpha^2 0(\sigma) + 0(\sigma^2), \\
&= \alpha^2 0(\sigma) + 0(\sigma^2).
\end{aligned}
$$

Hence there exists a mapping

$$
\begin{aligned}
I\!R^+ &\rightarrow I\!R^n \\
\sigma &\rightarrow r(\sigma)
\end{aligned}
$$

with $r(\sigma) = \alpha^2 0(\sigma) + 0(\sigma^2)$ such that the path $x(\sigma) = y(\sigma) + r(\sigma)$ satisfies

$$
g_J(x(\sigma)) = 0,
$$

For $i \in \bar{I}$ we similarly obtain

$$
g_i(x(\sigma)) \leq \sigma(u_i - \gamma) + \alpha^2 0(\sigma) + 0(\sigma^2)
$$

If α is small enough, $-\sigma\gamma + \alpha^2 0(\sigma) < 0$. In that case $x(\sigma)$ is, for $\alpha > 0$ small enough, feasible for $P(\sigma u)$ and

$$
\frac{\inf P(\sigma u) - f(\bar{x})}{\sigma^{1/2}} \leq \frac{f(x(\sigma)) - f(\bar{x})}{\sigma^{1/2}} \rightarrow -\alpha.
$$

This proves the Theorem. $\square$

We now consider the weak second-order sufficiency condition. For this we have to define the critical cone (at $\bar{x}$ for problem $P(0)$) :

$$
C = \left\{ d \in I\!R^n; \nabla f(\bar{x})^t d \leq 0; g'(\bar{x})d \overset{I}{<<} 0 \right\},
$$

and the following mapping

$$
H(x, \lambda) := \lambda_0 \nabla^2 f(x) + \sum_{i=1}^{p} \lambda_i \nabla^2 g_i(x).
$$

The weak second order sufficiency conditions is as follows :

$$
(3.1) \qquad \left\{ \begin{array}{l} \text{For all } d \text{ in } C - \{0\} \text{ there exists some multiplier } \lambda \\ \text{associated to } \bar{x} \text{ such that } d^t H(\bar{x}, \lambda) d > 0, \end{array} \right.
$$

Theorem 3.2 *If hypotheses $H1, H2, H3$ and (3.1) hold, and $\Lambda_1 = \emptyset$, then any solution $x(\sigma)$ of problem $P(\sigma u)$ satisfies*

$$
x(\sigma) = \bar{x} + 0(\sigma^{1/2}).
$$

Proof Let $x(\sigma)$ be a solution of $P(\sigma u)$. We may always write $x(\sigma) = \bar{x} + \alpha(\sigma)d(\sigma)$ with $\alpha(\sigma) \in \mathbb{R}^+$ and $\|d(\sigma)\| = 1$. If, for a given value of $\sigma, \alpha(\sigma) = 0$ the desired estimate is obtained for that value of σ, otherwise we have

$$\begin{aligned} g(x(\sigma)) &= g(\bar{x}) + \alpha(\sigma)g'(\bar{x})d(\sigma) + \frac{\alpha(\sigma)^2}{2}d(\sigma)^t\nabla^2 g(\bar{x})d(\sigma) + 0(\alpha(\sigma)^3) \\ & \overset{I}{<<} \sigma u_i \end{aligned}$$

Let λ be a degenerate multiplier associated to $\bar{x}$. Multiplying the above equalities or inequalities by λ_i and summing over i we obtain (as $\lambda_I \geq 0$) :

$$\frac{\alpha(\sigma)^2}{2}d(\sigma)^t H(\bar{x},\lambda)d(\sigma) + 0(\alpha(\sigma)^3) \leq \sigma\lambda^t u,$$

hence if $\alpha(\sigma) \neq 0$

$$d(\sigma)^t H(\bar{x},\lambda)d(\sigma) \leq \frac{2\sigma}{\alpha(\sigma)^2}\lambda^t u + 0(\alpha(\sigma)).$$

If the conclusion does not hold we have for some d, limit-point of $d(\sigma)$ (hence $\|d\| = 1$) when $\sigma \searrow 0$:

(3.2) $$d^t H(\bar{x},\lambda)d \leq 0 \text{ for all } \lambda \text{ in } \Lambda_0.$$

But from

$$g(x(\sigma)) = \alpha(\sigma)g'(\bar{x})d(\sigma) + 0(\alpha(\sigma)) \overset{I}{<<} \sigma u$$

we deduce that

$$g'(\bar{x})d \overset{I}{<<} 0.$$

By Theorem 3.1, for σ small enough :

$$0 \geq \frac{f(x(\sigma)) - f(\bar{x})}{\alpha(\sigma)} = \nabla f(\bar{x})^t d(\sigma) + 0(\alpha(\sigma)),$$

hence $\nabla f(\bar{x})^t d \leq 0$: this implies that d is in C. As $d \neq 0$ and $\|d\| = 1$ inequality (3.2) is in contradiction with (3.1). This proves the theorem. □

From Theorem 3.1. and Theorem 3.2. we deduce the

Corollary 3.1 *Under the hypotheses of Theorem 3.2. there exists* $m > 0, M > 0$ *such that the solutions* $\bar{x}(\sigma u)$ *of* $P(\sigma u)$ *satisfy for* $\sigma > 0$ *small enough*

$$m\sigma^{1/2} \leq \|\bar{x}(\sigma u) - \bar{x}\| \leq M\sigma^{1/2}.$$

We now try to characterize the limit points of $\sigma^{-1/2}(\bar{x}(\sigma u) - \bar{x})$, with $\bar{x}(\sigma u) \in S(\sigma u)$. Let d^0 be such a limit point. Expanding f and g up to the first order and using Theorem 3.1, we find that d^0 is a critical direction, as observed in [6]. Also, normalizing if necessary the multipliers associated to $\bar{x}(\sigma u)$ we find that (for the extracted sequence $\{\sigma^k\}$ corresponding to d^0) they have a limit-point λ^0 in Λ_0. If $\lambda_i^0 \neq 0$ the constraint i is active for σ^k small, hence $g_i'(\bar{x})d^0 = 0$. In summary, d^0 is an element of the following set :

$$C^0 = \{d \in C; \exists \lambda \in \Lambda_0; \lambda_i > 0 \Rightarrow g_i'(\bar{x})d^0 = 0\}.$$

For any d in C^0 we denote

$$I(d) = \{i \in \bar{I}; g_i'(\bar{x})d = 0\}.$$

We now consider the following subproblem

$$SP(u) \begin{cases} \min_{d^0,d^1} \nabla f(\bar{x})^t d^0; d^0 \in C^0; d^1 \in \mathbb{R}^n; \\ g'(\bar{x})d^1 + \frac{1}{2}(d^0)^t \nabla^2 g(\bar{x}) d^0 \overset{I(d)}{<<} u. \end{cases}$$

Here $d^t \nabla^2 g(\bar{x}) d$ is a vector whose i^{th} component is $d^t \nabla g_i(\bar{x}) d$. We actually prove that any limit point d^0 of $\sigma^{-1/2}(\bar{x}(\sigma u) - \bar{x})$ is associated to some d^1 such that (d^0, d^1) is a solution of $SP(u)$. There may be some "parasitic" solutions to $SP(u)$ (i.e. solutions that do not correspond to a limit point of the type $\sigma^{-1/2}(\bar{x}(\sigma u) - \bar{x})$) but we prove that any such parasitic solution gives rise to a suboptimal path.

Theorem 3.3 *Under the hypothesis of Theorem 3.2, the following holds :*

(i) *Any d^0, limit point of $\sigma^{-1/2}(\bar{x}(\sigma u) - \bar{x})$, with $\bar{x}(\sigma u) \in S(\sigma u)$, is associated to some d^1 in $\mathbb{R}^n$ such that (d^0, d^1) is a solution of $SP(u)$.*

(ii) *To each feasible solution d^0, d^1 in $SP(u)$ is associated a path $x(\sigma) = \bar{x} + \sigma^{1/2} d^o + o(\sigma)^{1/2})$ with $x(\sigma) \in F(\sigma u)$. If in addition (d^0, d^1) is a solution of $SP(u)$ then*

$$f(x(\sigma)) = \inf P(\sigma u) + o(\sigma^{1/2}).$$

Proof

a) Let d^0 be the limit of $(\sigma^k)^{-1/2}(x^k - \bar{x})$ with $\sigma^k \searrow 0$ and $x^k \in S(\sigma^k u)$. We already know that $d^0 \in C^0$. We may write

$$x^k = \bar{x} + (\sigma^k)^{1/2}(d^0 + \delta^k),$$

with $\delta^k \in \mathbb{R}^n, \delta^k \to 0$. Expanding the constraints up to the first order in σ we get :

$$g_i(x^k) = (\sigma^k)^{1/2} g_i'(\bar{x})\delta^k + \frac{\sigma^k}{2}(d^0)^t \nabla^2 g_i(\bar{x}) d^0 + o(\sigma^k) \overset{I(d^0)}{<<} \sigma u_i,$$

hence

$$g_i'(\bar{x})\frac{\delta^k}{(\sigma^k)^{1/2}} + \frac{1}{2}(d^0)^t\nabla^2 g_i(\bar{x})d^0 + \frac{o(\sigma^k)}{\sigma^k} \overset{I(d^0)}{<<} u_i$$

But the set

$$\left\{v; \exists d^1; g'(\bar{x})d^1 \overset{I(d^0)}{<<} v\right\}$$

is closed (see Bonnans-Launay [2], part 4) therefore $u - \frac{1}{2}(d^0)^t\nabla^2 g(\bar{x})d^0$ is in this set. We conclude that there exists $d^1 \in I\!R^n$ such that (d^0, d^1) is feasible for $SP(u)$.

b) We now prove the first statement of (ii), and the end the proof. For this purpose we consider the path

$$x(\sigma) = \bar{x} + (\sigma - \sigma^{1.3})^{1/2}d^0 + (\sigma - \sigma^{1.3})d^1 + \sigma^{1.3}\bar{d} + r(\sigma),$$

where $\bar{d}$ is again given by Lemma 2.1 and $r(\sigma)$ is a correction term. If $r(\sigma) = 0$ we get, using the fact that (d^0, d^1) is feasible for $SP(u)$:

$$\begin{aligned} g_J(x(\sigma)) &= (\sigma - \sigma^{1.3})[g_J'(\bar{x})d^1 + \frac{1}{2}(d^0)^t\nabla^2 g_J(\bar{x})d^0] \\ &\qquad + \sigma^{1.3}g_J'(\bar{x})\bar{d} + 0(\sigma^{3/2}), \\ &= \sigma u_J + 0(\sigma^{3/2}). \end{aligned}$$

As $\{g_i(\bar{x})\}_{i\in J}$ is linearly independent we may find $r(\sigma) = 0(\sigma^{3/2})$ such that $g_J(x(\sigma)) = \sigma u_J$. Expanding $g_i(x(\sigma))$ as before, for $i \in I(d)$ we find

$$g_i(x(\sigma)) \leq \sigma u_i - \sigma^{1.3}\gamma + 0(\sigma^{3/2}),$$

hence $g_i(x(\sigma)) < \sigma u_i$ for σ small. If $i \in \bar{I} - I(d)$ then $g_i'(\bar{x})d^0 < 0$ so that $g_i(x(\sigma)) < \sigma u_i$ is also satisfied. We have proved that $x(\sigma) \in F(\sigma u)$ when σ is small. In addition we obviously have

$$f(x(\sigma)) = f(\bar{x}) + \sigma^{1/2}\nabla f(\bar{x})^t d^0 + o(\sigma^{1/2}).$$

This, Theorem 3.1 and point (a) imply that $\nabla f(\bar{x})^t d^0$ must be minimum when d^0 is as in point (i). Conversely if (d^0, d^1) is any solution of $SP(u)$ we deduce the last statement of Theorem 3.3 with the above expansion of $f(x(\sigma))$. □

Remark 3.1

(i) From Theorem 3.3 we deduce that when $\sigma \searrow 0, \sigma^{-1/2}(\inf P(\sigma u) - \inf(P)) \to \inf(SP(u))$.

(ii) The proof of Theorem 3.3 gives us a means to compute effectively (e.g. numerically) a suboptimal path.

(iii) Assume that no second-order sufficiency condition holds. Then $\inf SP(u)$ may be $-\infty$. Considering the paths associated to the feasible solutions of SP(u) (whose existence does not depend on the second-order conditions) we deduce that in this case

$$\limsup_{\sigma\searrow 0} \frac{\inf P(\sigma u) - \inf P(0)}{\sigma^{1/2}} = -\infty. \square$$

4 THE CASE $\Lambda_1 \neq \emptyset$

We assume in this section that $\Lambda_1 \neq \emptyset$. Let us define

$$\Lambda^*(u) = \{\lambda \in \Lambda_1; -\lambda^t u = \max\{-\lambda^t u, \lambda \in \Lambda_1\}\}.$$

As observed in Gauvin-Janin [4] Hypotheses **(H3)** is equivalent to the requirement that $\Lambda^*(u)$ is non empty and bounded. We will use the strengthened second-order sufficiency condition :

(4.1) For any d in $C - \{0\}$ there exists λ in $\Lambda^*(u)$ such that $d^t H(\bar{x}, \lambda) d > 0$.

The stability of the perturbed solution under such hypothesis has been already obtained

Theorem 4.1 (Gauvin-Janin, [4]). *If $\Lambda_1 \neq \emptyset$, hypotheses (H1), (H2), (H3) and (4.1) imply that any solution $x(\sigma)$ of $P(\sigma u)$ satisfies*

$$\limsup_{\sigma\searrow 0} \|x(\sigma u) - \bar{x}\|/\sigma < +\infty. \square$$

Our goal is now to compute the directional derivatives of the solutions. We quote the following result.

Lemma 4.1 *Under the hypothesis of Theorem 4.1 the following holds :*

(i) *If K is a limit-point of the set of active contraints of the solutions of $P(\sigma u)$ when $\sigma \searrow 0$ then $\bar{x}$ satisfies the MF hypotheses for the set K.*

(ii) *The set of Lagrange multipliers associated to $x(\sigma u)$ remains bounded and its limit-points are in $\Lambda^*(u)$.*

(iii) *For $\sigma > 0$ small enough one has*

$$f(x(\sigma u)) = \max\{L(x(\sigma u), \lambda) - \sigma \lambda^t u; \lambda \in \Lambda^*(u)\}.$$

Proof Let $\bar{d}$ be given by Lemma 2.1. Define

$$w(\sigma) := \bar{d} - (x(\sigma u) - \bar{x})/\sigma,$$

If $x(\sigma)$ has K as a set of active constraints then for all i in K by Theorem 4.1 : $\sigma u_i = g_i(x(\sigma u)) = g_i'(\bar{x})(x(\sigma u) - \bar{x}) + o(\sigma)$ hence

$$\begin{aligned} g_i'(\bar{x})w(\sigma) &= g_i'(\bar{x})\bar{d} - (\sigma u_i + o(\sigma))/\sigma \\ &\leq -\gamma + o(\sigma)/\sigma. \end{aligned}$$

We may consider a sequence $\sigma^k \searrow 0$. Then $\{w(\sigma^k)\}$ is bounded by Theorem 4.1. Any limit-point w satisfies $g_i'(\bar{x})w \leq -\gamma, i \in K$, and similarly we can prove that $g_J'(\bar{x})w = 0$. This proves (i).

From (i) and the fact that the MF condition is stable by perturbation we deduce that the set of Lagrange multipliers associated to $x(\sigma u)$ is non empty and uniformly bounded when $\sigma \searrow 0$. If $\lambda(\sigma)$ is such a multiplier we have $\lambda_i(\sigma)(g_i(x)(\sigma u) - \sigma u_i) = 0$ for all i in I, hence if $\bar{\lambda}$ is a limit-point of $\{\lambda(\sigma)\}$ we also have (discussing the different cases) :

$$\bar{\lambda}_i(g_i(x(\sigma u)) - \sigma u_i) = 0.$$

Consequently for this Lagrange multiplier :

$$f(x(\sigma u)) = L(x(\sigma u), \bar{\lambda}) - \sigma\bar{\lambda}^t u. \tag{4.2}$$

On the other hand if λ is a Lagrange multiplier, from $\lambda_I \geq 0$ we deduce $\lambda_i(g_i(x(\sigma u) - \sigma u) \leq 0$ for all i in $I \cup J$ hence

$$f(x(\sigma u)) \geq L(x(\sigma u), \lambda) - \sigma\lambda^t u \ \ \forall\lambda \in \Lambda_1.$$

This and (4.2) prove

$$f(x(\sigma u), \lambda) = \max\{L(x, \lambda) - \sigma\lambda^t u; \lambda \in \Lambda_1\}.$$

Hence to prove (ii) and (iii) we just have to prove that $\bar{\lambda}$ is in $\Lambda^*(u)$.

From Theorem 4.1, (4.2) and the first -order optimality system of (P) we deduce

$$f(x(\sigma u)) = f(\bar{x}) - \sigma\bar{\lambda}^t u + 0(\sigma^2),$$

hence for the values of σ associated to K

$$\lim_{\sigma\searrow 0}(f(x(\sigma u) - f(\bar{x}))/\sigma = -\bar{\lambda}u$$

by Theorem 4.1. Notice that $(x(\sigma u) - \bar{x})/\sigma$ has at least a limit-point $\tilde{d}$ which is a feasible point of the linear program

$$\min_d \nabla f(\bar{x})^t d \quad ; \quad g'(\bar{x})d \overset{I}{<<} u. \tag{4.3}$$

But the dual of (4.3) is

$$\max -\lambda^t u \quad ; \quad \lambda \in \Lambda_1.$$

The relation $\frac{d}{d\sigma}f(x(\sigma u)) = \nabla f(\bar{x})^t, \tilde{d} = -\bar{\lambda}^t u$ hence implies that $\tilde{d}$ is a solution of (4.3) and that $\bar{\lambda}$ is a solution of its dual, i.e. $\bar{\lambda}$ is in $\Lambda^*(u)$. $\square$

The above relations will allow us to expand $f(x(\sigma u))$ up to the second order. Indeed let us define

$$Q^*(d) := \max\{d^t H(\bar{x}, \lambda)d; \lambda \in \Lambda^*(u)\}.$$

Then

Proposition 4.1 *Under the hypotheses of Theorem 4.1, let $\{\sigma^k\}$ be such that $\sigma^k \searrow 0$ and that x^k is a solution of $P(\sigma^k u)$ and $(x^k - \bar{x})/\sigma^k \to d$. Then*

$$F(x^k) = f(\bar{x}) - \sigma^k \lambda^t u + \frac{(\sigma^k)^2}{2} Q^*(d) + o((\sigma^k)^2).\square$$

<u>Proof</u> From Theorem 4.1, Lemma 4.1 (iii) and the boundedness of $\Lambda^*(u)$ we get for any $\bar{\lambda}$ in $\Lambda^*(u)$:

$$F(x^k) = \max\{f(\bar{x}) + \frac{1}{2}(x^k - \bar{x})^t H(\bar{x}, \lambda)(x^k - \bar{x}), \lambda \in \Lambda^*(u)\} - \sigma^k \bar{\lambda}^t u + o((\sigma^k)^2)$$

with $x^k - \bar{x} = \sigma^k d + o(\sigma^k)$ we get the result. $\square$

We now try to characterize the limit-points of $\sigma^{-1}(\bar{x}(\sigma u) - \bar{x})$ when $\sigma \searrow o$, with $\bar{x}(\sigma u) \in S(\sigma u)$. We consider the following sets :

$$\begin{aligned} F^*(u) &:= \{x \in F(u); \exists \lambda \in \Lambda^*(u); \lambda_i > 0 \Rightarrow g_i(x) = u_i, i \in I\}, \\ C^*(u) &:= \{d \in \mathbb{R}^n; g'(x)d << u, \nabla f(\bar{x})^t d \to \min\}. \end{aligned}$$

By Lemma 4.1 we have $S(\sigma u) \subset F^*(\sigma u)$ for σ small. The set $C^*(u)$ is the set of solutions of a linear program and can be characterized using the associated multipliers :

$$C^*(u) = \{d \in \mathbb{R}^n; g'(x)d << u; \exists \lambda \in \Lambda^*(u); \lambda_i > 0 \Rightarrow g_i'(\bar{x})d = u_i, i \in I\}$$

Here we see that $C^*(u)$ is a candidate for the "linearization" of the set-valued mapping $u \to F^*(u)$ at $u = 0$, and at point $\bar{x}$. We will use the following directional qualification property :

$$DQ^*(u) \quad \forall d \in C^*(u); \exists x(\sigma u) = \bar{x} + \sigma d + o(\sigma), x(\sigma u) \in F^*(\sigma u).$$

Theorem 4.2 *Under the hypotheses of Theorem 4.1 and $DQ^*(u)$ then :*

(i) *Let $\tilde{d}$ be a limit-point of $(\sigma)^{-1}(\bar{x}(\sigma u) - \bar{x})$ with $\bar{x}(\sigma u) \in S(\sigma u)$. Then $\tilde{d}$ is a solution of the "pseudo-quadratic" problem*

$$\min Q^*(d); d \in C^*(u). \tag{4.4}$$

(ii) *Conversely to each solution $\tilde{d}$ of this problem is associated a path $x(\sigma u) = \bar{x} + \sigma\tilde{d} + o(\sigma)$ with $x(\sigma u) \in F^*(\sigma u)$ and*

$$f(x(\sigma u)) = \inf P(\sigma u) + o((\sigma)^2).\square$$

Proof Let $\tilde{d}$ be as in point (i). Then $\tilde{d}$ in in $C(u)$. From the proof of Lemma 4.1 we get $\nabla f(\bar{x})^t\tilde{d} = -\bar{\lambda}^t u$ for any $\bar{\lambda}$ in $\Lambda^*(u)$. This implies that $\tilde{d}$ is in $C^*(u)$, hence is feasible for (4.4).

Let d be in $C^*(u)$ and $x(\sigma)$ the path associated by $DQ^*(u)$. Expanding $f(x(\sigma))$ as in Proposition 4.1 we get (here again $\bar{\lambda}$ is any element of $\Lambda^*(u)$) :

$$f(x(\sigma)) = f(\bar{x}) - \sigma\bar{\lambda}^t u + \frac{(\sigma)^2}{2}Q^*(d) + o((\sigma)^2).$$

With Proposition 4.1 this implies that $Q^*(.)$ must be minimum at $\tilde{d}$. The rest of the Theorem follows. $\square$

Remark 4.1 Assume that $\Lambda^*(u)$ reduces to a singleton $\{\lambda\}$ which will be the case of for almost all directions u (for the Lebesgue measure). Applying the characterization of Kyparisis [9] to the problem

$$\min \nabla f(\bar{x})^t d; d \in C(u)$$

and defining $K := J \cup \{i \in I; \bar{\lambda}_i > 0\}$, we see that d^0 being in $C^*(u)$ the following is satisfied :

(i) $\{\nabla g_i(\bar{x}), i \in K\}$ is linearly independant.

(ii) $\exists d^\sharp \in \ker g'_K(\bar{x}); g'_i(\bar{x})d^\sharp < 0, \forall i \in I(d) - K$.

From this it can be easily deduced that $CQ^*(u)$ holds, hence we may apply Theorem 4.2. Also our Theorem 4.2 implies Theorem 5.1 of Gauvin-Janin [4] in which the derivative of the solution is obtained under stronger hypothesis.

Remark 4.2 We may compare Theorem 4.2 with corresponding results of Shapiro [11]. An improvement is that Q^* is defined by a maximization over $\Lambda^*(u)$ and not all Λ_1. Consequently, as $\Lambda^*(u)$ reduces to a singleton a.e. in u, problem (4.5) is a true quadratic problem, a.e. in u.

5 EXAMPLES

We give two two-dimensional examples for which the hypothesis of Mangasarian and Fromovitz is not satisfied, the first with $\Lambda_1 = \emptyset$. We show how our theorems allow to compute effectively suboptimal paths on these examples.

First example The problem is

$$\min x_1; \quad x_1^2 - x_2 \le u_1, \quad x_1^2 + x_2 \le u_2.$$

One has $F(0) = \{0\}$ hence the original problem has the unique solution $\bar{x} = 0$ and $H2$ is easy to check. We easily obtain $\Lambda_1 = \emptyset$ and

$$\Lambda_0 = \left\{ \begin{pmatrix} \alpha \\ \alpha \end{pmatrix}; \alpha \in I\!R^{+*} \right\}.$$

Consequently $H3$ will be satisfied iff $u_1 + u_2 > 0$. We assume that this is the case. We also obtain

$$C = \{d \in I\!R^2; d_1 \le 0, d_2 = 0\},$$

$$\frac{1}{2} d^t H(\bar{x}, \lambda) d = 2\alpha d_1^2,$$

with $\lambda = (\alpha \quad \alpha)^t, \alpha \ge 0$. This implies (3.1). We may apply Theorem 3.3. Here $C^0 = C$ hence $SP(u)$ reduces to

$$\min_{d^0, d^1} d_1^0; \quad d_1^0 \le 0; \quad d_2^0 = 0 \quad ; \quad (d_1^0)^2 - d_2^1 \le u_1,$$
$$(d_1^0)^2 + d_2^1 \le u_2.$$

The two last inequalities are equivalent to

$$(d_1^0)^2 \le \min(u_1 + d_2^1, u_2 - d_2^1).$$

In order to get d_1^0 minimum we have to maximize the right hand side with respect to d_2. The maximum is obtained when $d_2^1 = \frac{1}{2}(u_2 - u_1)$($d_1^1$ having any value) and then $d_1^0 = -\sqrt{\frac{u_2+u_1}{2}}$ is optimal. Hence there is a unique component d^0 for the solutions of $SP(u)$. Consequently the solutions $\bar{x}(\sigma u)$ of the perturbed problem satisfy

$$\bar{x}(\sigma u) = \bar{x} - \sigma^{1/2} \begin{pmatrix} \sqrt{\frac{u_1+u_2}{2}} \\ 0 \end{pmatrix} + o(\sigma^{1/2}).$$

We may check this result by computing the exact solution of the perturbed problem. They saturate the two constraints, hence there is a unique solution

$$\bar{x}(\sigma u) = \begin{pmatrix} -\sigma^{1/2}\sqrt{\frac{u_1+u_2}{2}} \\ \sigma\frac{u_2-u_1}{2} \end{pmatrix}.$$

Second example The problem is

$$\min x_2; \quad x_1^2 - x_2 \le u_1, \quad x_1^2 + x_2 \le u_2.$$

We only changed the criterion. Again $\bar{x} = 0$, $H1$ and $H2$ are satisfied, Λ_0 is as before and $H3$ is satisfied iff $u_1 + u_2 > 0$ which is assumed to hold. Now

$$\begin{aligned} \Lambda_1 &= \left\{ \lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}; \lambda_1 - \lambda_2 = 1, \lambda_1 \geq 0, \lambda_2 \geq 0 \right\}, \\ &= \left\{ \lambda = \begin{pmatrix} 1 + \lambda_2 \\ \lambda_2 \end{pmatrix}; \lambda_2 \geq 0 \right\}, \end{aligned}$$

and $\Lambda^*(u)$ is the set of element of the form $(1 + \lambda_2, \lambda_2)^t$ for which λ_2 is solution of

$$\min(1 + \lambda_2)u_1 + \lambda_2 u_2; \quad \lambda_2 \geq 0.$$

As $u_1 + u_2 > 0$ we get $\Lambda^*(u) = \{\bar{\lambda}\}$ with $\bar{\lambda} = (1 \quad 0)^t$, hence $CQ^*(u)$ holds. Also

$$C = \{d \in I\!R^2; d_2 = 0\}$$

and

$$\frac{1}{2} d^t H(\bar{x}, \bar{\lambda}) d = d_1^2,$$

hence (4.1) holds. We may apply Theorem 4.1. The quadratic problem to be considered is

$$\min d_1^2; \quad -d_2 = u_1, \quad d_2 \leq u_2.$$

This problem has unique solution: $(0 \quad -u_1)^t$. We have proved that any optimal solution $\bar{x}(\sigma u)$ of $P(\sigma u)$ satisfies

$$\bar{x}(\sigma u) = \bar{x} + \sigma \begin{pmatrix} 0 \\ -u_1 \end{pmatrix} + o(\sigma).$$

In fact the exact solution of $P(\sigma u)$ is just $(0 \quad -\sigma u_1)^t$.

References

[1] Bonnans J.F., *A semi-strong sufficiency condition for optimality in nonconvex programming and its connection to the perturbation problem*, J. Optim. Th. Appl. 60(1989), 7-18.

[2] Bonnans J.F., Launay G., *On the stability of sets defined by a finite number of equalities and inequalities*, to appear.

[3] Fiacco A.V., *Introduction to sensitivity and stability analysis in nonlinear programming*, Academic Press, New York, 1983.

[4] Gauvin J., Janin R., *Directional behaviour of optimal solutions in nonlinear mathematical programming*, Math. Oper. Res. 13 (1988), 629-649 .

[5] Gauvin J., Janin R., *Directional derivative for the value function in parametric optimization* (1988), to appear.

[6] Gauvin J., Janin R., *Directional lipschitzian optimal solutions and directional derivative for the optimal value function in nonlinear mathematical programming*, to appear.

[7] Gollan B., *On the marginal function in nonlinear programming*, Math. Oper. Res. 9 (1984), 208-221.

[8] Jittorntrum K. *Solution point differentiability without strict complementarity in nonlinear programming*, Math. Programming Study 21 (1984) 127-138.

[9] Kyparisis J., *On uniqueness of Kuhn-Tucker multipliers in nonlinear programming*, Math. Programming 32 (1985), 242-246.

[10] Mangasarian O.L., Fromovitz S., *The Fritz John necessary optimality conditions in the presence of equality and inequality constraints*, J. Math. Anal. Appl. 7(1967), 37-47.

[11] Shapiro A. *Sensitivity analysis of nonlinear programs and differentiability properties of metric projections*, SIAM J. Control Optimiz. 26 (1988), 628-645.

Imprimé en France
par
l'Institut National de Recherche en Informatique et en Automatique